Available online at www.sciencedirect.com
ScienceDirect
Computers and Electronics in Agriculture 53 (2006) 98–112
Computers and electronics in agriculture
www.elsevier.com/locate/compag

Directed sampling using remote sensing with a response surface sampling design for site-specific agriculture☆

Glenn J. Fitzgerald[a,*], Scott M. Lesch[b], Edward M. Barnes[c], William E. Luckett[a]

[a] *U.S. Arid Lands Agricultural Research Center, 21881 N. Cardon Lane, Maricopa, AZ 85239, USA*
[b] *USDA-ARS George E. Brown, Jr. Salinity Lab, 450 West Big Springs Road, Riverside, CA 92507, USA*
[c] *Cotton Incorporated, 6399 Weston Parkway, Cary, NC 27513, USA*

Received 9 January 2006; received in revised form 18 April 2006; accepted 25 April 2006

Abstract

Remotely sensed imagery provides contiguous spatial coverage of a field and can be used as a surrogate to measure crop and soil attributes. Empirical regression models are often used to convert imagery to attribute maps, when an *a priori* linear relationship can be assumed to exist between the imagery and ground attributes. In this study, we used the response surface approach incorporated in the EC_e Sampling, Assessment, and Prediction (ESAP) software to create ground sampling designs from input imagery in order to develop regression equations for predicting crop height and width attributes in a 3.4-ha cotton field. We examined both the reliability of this model-based sampling approach as well as the validity of the assumed linear models using multiple-date imagery and sample data collected from a 3-year remote sensing experiment. Predictions of height and width from regressions between the imagery and ground sampling at the calibration locations gave coefficients of determination for height ranging from 0.34 to 0.90 and for width, 0.30 to 0.94. All regression models but one were statistically significant at the $\alpha = 0.01$ level. To test the reliability of the sampling approach, the regression models developed during the first year were used to predict additional crop height and width attributes at a randomly chosen set of validation sites. Multiple statistical tests indicated that these predictions were both unbiased and within the specified precision of the estimated regression equations. This regression-based directed sampling and estimation method requires fewer points than co-kriging to develop reliable imagery-crop attribute relationships, and thus is potentially less expensive. We hypothesize that other variables such as crop nitrogen might also be accurately predicted using this approach as long as the crop attribute and spectral index meet the model assumptions. Maps of crop attributes and/or soil properties could be used by farmer consultants to schedule variable-rate applications of chemicals or as inputs to crop simulation models providing a spatial extension to their time-series nature.
Published by Elsevier B.V.

Keywords: Directed sampling; Remote sensing; Precision agriculture; Response surface sampling; Spatial autocorrelation; Cotton

1. Introduction

One limitation to the application of remotely sensed imagery in site-specific agriculture is the lack of cost-effective sampling procedures to convert imagery to maps of the primary crop or soil attribute of interest. Depending on

☆ Mention of trade names or proprietary products in this paper does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

* Corresponding author. Tel.: +1 520 316 6375 (USA)/+61 3 5362 2145 (Australia); fax: +1 520 316 6330.
E-mail addresses: gfitzgerald@uswcl.ars.ag.gov, glenn.fitzgerald.@dpi.vic.gov.au (G.J. Fitzgerald).

0168-1699/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.compag.2006.04.003

the number of samples collected, ground sampling based on grids or management zones can be expensive and the placement of sampling locations within management zones can be ambiguous. Also, if not properly accounted for, spatial autocorrelation can lead to errors when developing relationships between imagery and ground data (Atkinson, 2002). Additionally, the scale of soil and plant sampling must correspond to imagery pixel size (Atkinson and Curran, 1995, 1997) and the remotely sensed index must be strongly correlated to the ground parameter of interest. Thus, to be statistically robust and cost-effective, an ideal sampling strategy using remotely sensed imagery should (i) account for factors of scale of variability, (ii) have available a spatially contiguous (covariate) data set of the field, (iii) reduce spatial autocorrelation as much as possible, and (iv) provide unambiguous geo-positioned sampling locations.

Directed (or "model-based") sampling uses the information contained in a covariate spatial data set (i.e., an image or fine grid of survey data) to select specified locations in a field for sampling (specific crop or soil attributes), based on the variability present in the covariate data set and the assumed form of the regression function (Lesch, 2005). The covariate data set can be two-dimensional (maps) of yield, topography, electromagnetic induction, soil texture, salinity or be image-derived, as discussed here. Directed sampling is advantageous when the general form of the regression function can be specified *a priori*, but a limited number of soil or crop samples must still be acquired to calibrate (i.e., estimate) the model parameters. In such situations, directed sampling techniques can typically be used to reduce the number of sampling locations needed for efficient parameter estimation.

The EC_e Sampling, Assessment, and Prediction-response surface sampling design (ESAP-RSSD) software program was used to direct all ground sampling performed in this study. This software generates spatially referenced sampling designs by selecting a minimum set of calibration samples based on the observed magnitudes and spatial locations of the data, with the explicit goal of optimizing the estimation of a multiple linear regression model (Lesch et al., 1995a, b, 2000; Lesch, 2005). The sampling algorithm also attempts to select calibration sampling locations that maximize the probability of generating spatially uncorrelated regression model residuals, thus allowing the analyst to use an ordinary regression model in place of a more complicated spatial linear model or geostatistical model (Schabenberger and Gotway, 2005). The software has been used extensively in salinity assessment and precision farming applications and was originally developed to facilitate the estimation of soil salinity (EC_e) from soil conductivity (EC_a) survey data (Rhoades et al., 1999; Corwin and Plant, 2005; Lesch et al., 2005). However, in principle, the underlying sampling methodology is broad enough to allow input of other types of geo-referenced survey data, such as imagery.

The objective of this research was to evaluate the ability of the ESAP-RSSD software to direct ground sampling by substituting aerial imagery for EC_a and crop height and width for EC_e, producing predictive maps of these crop attributes. The first year, both calibration and validation data were collected and subjected to various statistical analyses to test the effectiveness and reliability of the directed sampling approach. In all three years, the software was used to create directed sampling plans for estimating the regression equations that converted the aerial images into field maps of crop width and height.

2. Materials and methods

2.1. Description of field and location

The experimental site was a 3.4-ha field planted to cotton (*Gossypium hirsutum* L., cv. Delta Pine 448B) on 28 April 2001 and 15 April in 2002 and 2003 at the University of Arizona, Maricopa Agricultural Center (MAC) located approximately 40 km south of Phoenix (33°04′21″N; 111°58′45″W) at an elevation of 360 m above sea level (Fig. 1). The field straddles the transition between two soil series (Post et al., 1988): Mohall sandy loam (fine-loamy, mixed, hyperthermic Typic Haplargid) is dominant in the northeast portion of the field, and Casa Grande sandy clay loam (fine-loamy, mixed, hyperthermic Typic Natrargids) spans most of the southwest region. This is an arid area, receiving 185 mm of rainfall per year with maximum daily summer temperatures ranging from 25 to 46 °C. The field had an on-going tillage study imposed with two-row "skips" where no cotton was grown. The field was graded level and furrow irrigated from the east.

2.2. Remote sensing

In 2001, aerial imagery was acquired on three dates using an off-the-shelf Nikon Coolpix 950 (Nikon Inc., Torrance, CA) true-color three megapixel digital camera (Table 1). Flight elevations were 920 m (3000 ft) above ground level.

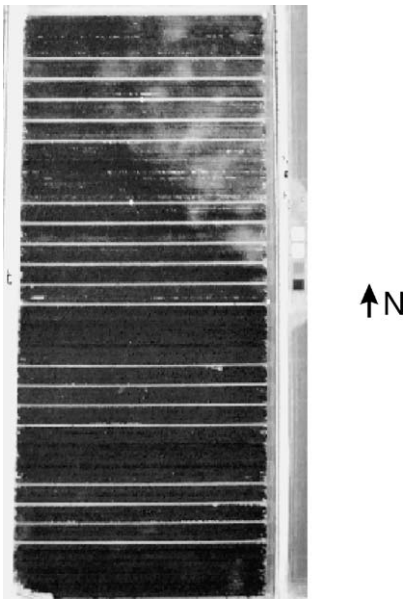

Fig. 1. Mid-season panchromatic image of cotton field used for directed sampling acquired with an off-the-shelf digital camera. Soil and roads appear bright, while dense cotton canopy appears dark. Four, 8 m by 8 m reference tarps used to convert 2002 and 2003 imagery to reflectance for NDVI calculations are visible along the eastern edge of the field. The bright east to west lines are 2-m wide "skip" rows where cotton was not planted allowing tractor access to the fields. The five east–west broad areas or "benches" were 40-row wide areas planted to cotton. A brighter soil with greater sand content is visible in the NE portion of the field (leading to poorer cotton stands and more visible soil).

The dates were chosen so that between 30 and 90% cover was established in the crop, which is a period during crop development when spectral measurements have more robust relations for estimating ground parameters. Very early in the season the proportion of bare soil is high, which masks the plant spectral signal (Huete, 1988; Jackson and Huete, 1991). Late in the season, after full cover, relationships change due to canopy closure (Carlson and Ripley, 1997) and senescence. Flights occurred between 1030 and 1100 h local time. The Nikon red band images were used to select

Table 1
Flight day of year (DOY), dates, field conditions, and image type

DOY (2001)	Date	Nominal field conditions	Image type
118	28 Apr 2001	Dry soil, planted Delta Pine 448B	–
198	17 Jul 2001	Dry soil, ~75% cover	Red uncal[a]
207	26 Jul 2001	Dry soil, ~80% cover	Red uncal
226	14 Aug 2001	Dry soil, ~95% cover	Red uncal
DOY (2002)			
105	16 Apr 2002	Dry soil, planted Delta Pine 448B	–
176	25 Jun 2002	Dry soil, ~45% cover	NDVI cal[b]
192	11 Jul 2002	Dry soil, ~75% cover	NDVI cal
207	26 Jul 2002	Irrigated 3 days before, ~85% cover	NDVI cal
DOY (2003)			
104	15 Apr 2003	Dry soil, planted Delta Pine 448B	–
168	17 Jun 2003	Dry soil, ~30% cover	NDVI cal[b]
176	25 Jun 2003	Irrigated 5 days before, ~45% cover	NDVI cal
217	5 Aug 2003	Irrigated 4 days before, ~85% cover	NDVI cal
237	25 Aug 2003	Dry soil, ~90% cover	NDVI cal

[a] Off-the-shelf wide-band digital camera, red band, uncalibrated.
[b] 3-CCD narrow-band camera, calibrated to reflectance using ground reference tarps.

the sampling locations. The red band was centered at 625 nm with a 100-nm bandwidth (full width half maximum). Field of view and pixel resolution varied depending on zoom setting but pixel resolutions were always less than 0.5 m.

Imagery was acquired on three dates in 2002 and four in 2003 using a Duncan MS3100 (Redlake Inc., San Diego, CA) camera that acquired three coincident, 8-bit images in three wavebands (Table 1). The wavebands were centered on 670, 720, and 790 nm with 10-nm bandwidths (full width half maximum). Flight elevation in both years was 1070 m (3500 ft) above ground level. The Duncan camera had a $15^\circ \times 20^\circ$ field of view resulting in a pixel resolution of about 0.4 m. Imagery on the dates reported here was acquired about 60–90 min after solar noon, between 1315 and 1400 local time.

Reference tarps (Group VIII Technologies, Provo, UT) with known nominal reflectance values of 4, 8, 48, and 64% were placed either adjacent to the field or in a neighboring field during each flight in 2002 and 2003, and images were acquired within 7 min of acquiring images of the tarps. Coincident, ground-based PS-II™ (Analytical Spectral Devices, Boulder, CO) radiometer measurements were made over the tarps near flight times to radiometrically calibrate the imagery to reflectance. The PS-II measures energy from 400 to 1100 nm. These procedures and calibration tarps are similar to those described by Moran et al. (2001). Geometric registration for all 3 years was accomplished by measuring coordinates of markers placed in the images with a Trimble Ag132 (Trimble Navigation Limited, Sunnyvale, CA) receiver with real-time differential correction. Repeated measures of the markers were made during the season to insure the locations were known to sub-meter accuracy.

2.3. Ground sampling

In 2001, there were 25 sampling sites: 12 calibration and 13 validation. The validation sites were 'fixed' in location for all dates while the calibration site locations varied due to the changing imagery information acquired from each flight. Imagery collected during the 1994 MADMAC experiments (Moran et al., 1996) was used to determine the initial 'fixed' validation sampling locations using normalized difference vegetation index (NDVI) values of the field on 14 June. On that date, the field contained a cotton crop that provided approximately 30% ground cover. The validation locations were chosen using a stratified-random sampling scheme, with the NDVI representing the stratification variable. Plant height and canopy width measurements were collected weekly at the validation locations and after each flight at the calibration locations, usually within 24 h. For 2001 and 2002, plant height and width were measured for five plants at each sampling site on each sample date at 0.5-m intervals. Canopy width was measured using a meter stick placed perpendicular to the row at each plant location. Rows were spaced 1.02 m (40 in.) apart. Width was a surrogate for percent cover. Plant samples were selected to represent the area around the sample location. In 2003, plant width was measured at three locations in a row and height of 10 plants was averaged.

2.4. Directed sampling and image processing

All image processing, including geo-registration, masking, classification, etc., was performed using ENVI software (Research Systems Inc., Boulder, CO). Raw imagery (digital numbers) was radiometrically calibrated to reflectance in 2002 and 2003 and converted to NDVI using the 670 and 790 nm wavebands from the Duncan camera. The NDVI was calculated as: $(790 - 670)/(790 + 670)$. Imagery from all dates was then geo-referenced with known ground points visible in the imagery. Since the intent was to provide locations to sample plants, the "skips" were masked out of the imagery using standard procedures in the ENVI software before entering data into the ESAP-RSSD software to avoid selecting non-plant sample sites in these areas (Fig. 2). In 2001 and 2002, image data were re-sampled to 2-m pixels so that the number of data points (pixels) was less than the software-imposed limit of 30,000 points and then input to ESAP-RSSD as text files containing coordinates and pixel data. In 2003, pixels were re-sampled to 1 m and the narrow strips of cotton were masked out of the imagery (Fig. 2g–i).

The ESAP-RSSD software produces sample designs that space the sample locations apart to minimize the possibility of spatial autocorrelation in the regression model residuals. Each sampling design is assigned an "optimization criterion" (OptCri), which is a measure of the uniformity of coverage of the selected sampling locations, normalized for the field size (Lesch et al., 2000). More specifically, for a sample of size n, the program calculates the approximate maximum possible separation distance (SD_p) that a uniformly spaced sampling pattern might achieve (assuming the field is rectangular). It then calculates the achieved average separation distance for the current design (SD_a) and computes

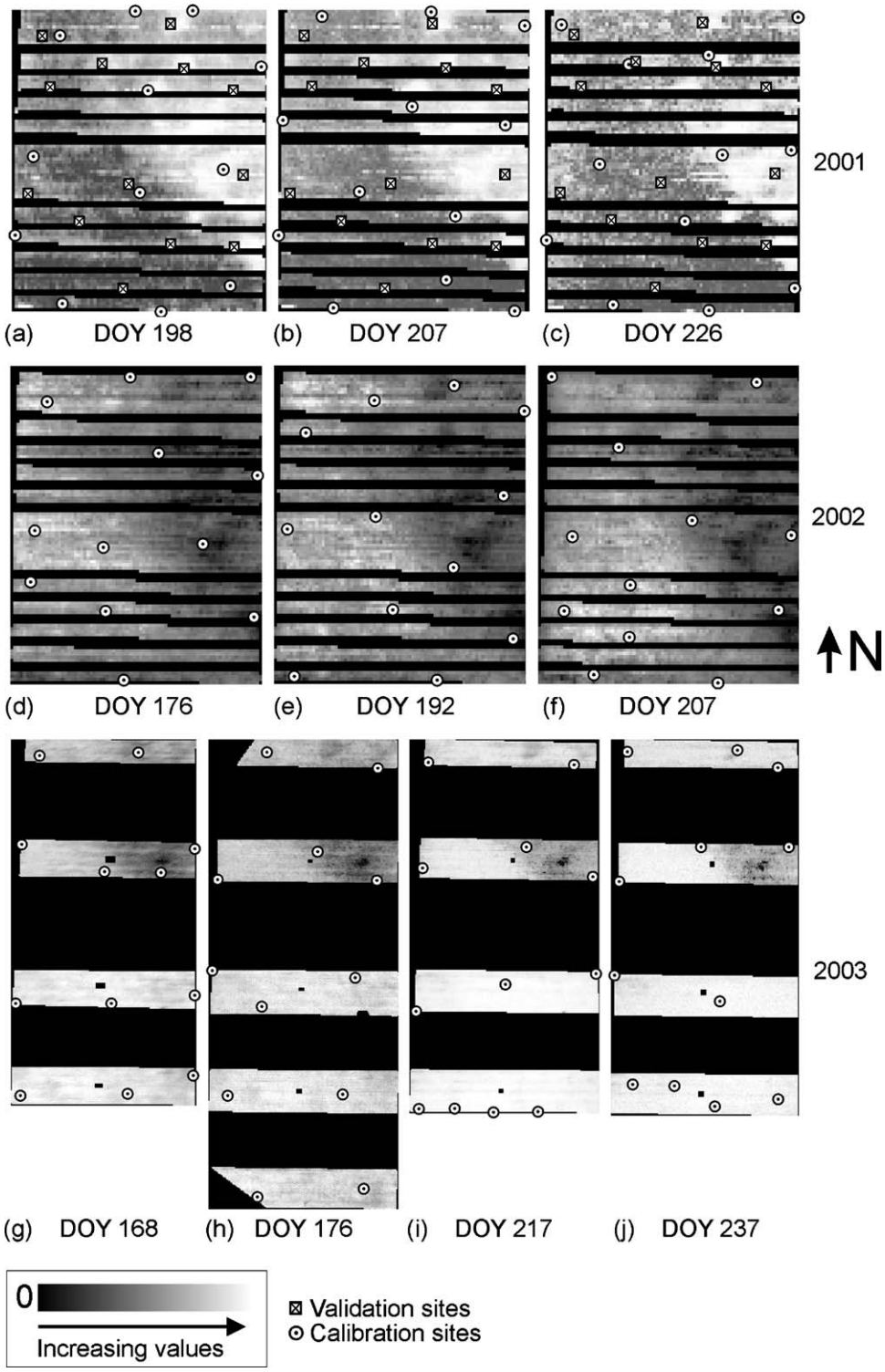

Fig. 2. Directed sampling locations for all 3 years and sampling dates. In 2001 (a–c) and 2002 (d–f) only the northern portion of the field was sampled, while in 2003 (g–j), only the wider "benches" were used for sampling and analysis (compare to Fig. 1). Validation sites were located in 2001 only. In 2001, the red band from an off-the-shelf digital camera was used to direct sampling. Higher values indicate sparser vegetation. In 2002 and 2003, NDVI was calculated and higher values indicate denser vegetation. Areas of skip rows where no cotton was planted were masked, as were field edges. In 2003, small squares in the middle of the benches were masked to exclude markers placed in the fields to facilitate image geo-registration.

the OptCri score as SD_p/SD_a. The operator can then choose a design based on the assigned OptCri. Typically, as was done here, the sample design with the lowest OptCri (best separation of points) is chosen. The default setting in the ESAP-RSSD program selects 12 geo-registered calibration locations for sampling, thus directing the user to these locations for ground sampling. Details of the selection procedure and use of the software can be found in Lesch et al. (2000). Outliers were removed from the input pixel values that were four or more standard deviations from the mean.

Sampling sites were located on the ground by entering as input the 12 calibration coordinates into a differential global positioning system and walking to them. After testing for statistical significance between the pixel values and ground data and residual spatial correlation (see Section 2.5), predictions of crop height and width were accomplished by fitting regression models between the image pixel values (red band or NDVI) and ground data at the sampling locations for each date. To test the validity of the sampling approach, the 13 fixed validation locations were used as independent data sets in 2001 to validate the crop attribute values predicted from the calibration sites. In 2002 and 2003, imagery was calibrated to predict height and width across the field without validation. Mean pixel values were extracted from the imagery in 3-pixel diameter areas around the sampling locations to develop the calibration regressions.

2.5. Statistical analyses

Six sequential steps were used during the modeling and validation of the 2001 data: (1) remove outliers in the regressions between imagery and ground data for each date, (2) check that the residual errors were normally distributed using distribution tests and residual plots for normality, (3) determine statistical significance of correlation models for the pooled data (calibration and validation sites), (4) produce regression statistics for pooled data, (5) perform a test for residual spatial independence (Moran test statistic, I_{MS}), and (6) perform F-tests to test the ability of the regression models to predict the ground values at the validation sites from the calibration data alone.

Of the 318 data points used in the analysis (10 dates × 12 calibration sites + 13 validation sites × 3 dates for each of two parameters), only two were considered outliers since the externally standardized residuals were four or more standard deviations from zero. After removing these outliers and testing for normality, coefficients of determination were calculated for pixel values (red or NDVI) versus crop attributes. Regression statistics were then calculated (Table 2) for those models that were statistically significant at the $\alpha = 0.01$ level. This included the Moran test statistic (I_{MS}) for spatial independence of residuals (Upton and Fingleton, 1985; Haining, 1990) and three separate F-tests for model validity, as described below.

In all of the analyses, a simple linear model was assumed to explain each crop attribute/pixel level relationship; i.e.,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{1}$$

where y represents the crop attribute, x represents the corresponding pixel level, (β_0, β_1) represent the unknown intercept and slope parameters, and ε represents the (normal) random error component. Note that this model can be conveniently expressed in matrix notation as

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \tag{2}$$

where all components are defined as before. After estimating Eq. (2) using the pooled data, the Moran test was calculated as

$$I_M = \frac{(\mathbf{e}'\mathbf{W}\mathbf{e})}{\mathbf{e}'\mathbf{e}} \quad \text{where } \mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta} \tag{3}$$

and $\mathbf{W}$ was defined to be a normalized inverse-distance squared weighting matrix. The corresponding normalized I_M test statistic (I_{MS}) was then defined as

$$I_{MS} = \frac{(I_M - E\{I_M\})}{\mathrm{Std}(I_M)} \tag{4}$$

where the expectation and variance of the test statistic were computed using the formulas given in Upton and Fingleton (1985).

Table 2
Regression statistics

Attribute	DOY	Date	Mean	Slope (S.E.)	*Y*-intercept	r^2	RMSE	I_{MS}
Height	198	17 Jul 2001	0.76 (0.026)	−0.0053 (0.0004)[a]	1.22 (0.39)	0.88	0.044	0.78
	207	26 Jul 2001	0.82 (0.020)	−0.0024 (0.0007)	1.07 (0.73)	0.34	0.084	0.31
	226	14 Aug 2001	0.89 (0.026)	−0.0080 (0.0016)[a]	1.38 (0.99)	0.55	0.077	0.90
Width	198	17 Jul 2001	0.75 (0.022)	−0.0044 (0.0006)	1.15 (0.57)	0.68	0.064	0.90
	207	26 Jul 2001	0.81 (0.017)	−0.0019 (0.0006)	1.00 (0.63)	0.30	0.072	0.09
	226	14 Aug 2001	0.97 (0.019)	n/a (truncated data)	–	–	–	–
Height	176	25 Jun 2002	0.46 (0.026)	0.62 (0.15)[b]	0.26 (0.51)	0.67	0.053	−0.74
	192	11 Jul 2002	0.70 (0.024)	0.34 (0.12)[c]	0.48 (0.80)	0.45	0.066	−0.01
	207	26 Jul 2002	0.84 (0.050)	1.0 (0.257)	0.01 (2.20)	0.60	0.012	0.07
Width	176	25 Jun 2002	0.47 (0.022)	0.49 (0.14)[b]	0.31 (0.50)	0.58	0.050	−0.71
	192	11 Jul 2002	0.74 (0.022)	–	–	0.01 ns	–	1.82[d]
	207	26 Jul 2002	0.87 (0.030)	0.56 (0.17)	0.40 (1.50)	0.51	0.077	0.17
Height	168	17 Jun 2003	0.31 (0.019)	0.79 (0.083)	0.19 (0.014)	0.90	0.022	1.94[e]
	176	25 Jun 2003	0.41 (0.028)	0.66 (0.114)	0.26 (0.028)	0.77	0.030	0.33
	217	5 Aug 2003	0.84 (0.035)	1.16 (0.102)	0.12 (0.064)	0.93	0.034	−1.02
	237	25 Aug 2003	0.89 (0.034)	1.22 (0.178)	0.12 (0.113)	0.83	0.052	0.37
Width	168	17 Jun 2003	0.32 (0.014)	0.48 (0.113)	0.24 (0.019)	0.65	0.030	0.74
	176	25 Jun 2003	0.47 (0.017)	0.61 (0.126)	0.33 (0.030)	0.70	0.033	0.15
	217	5 Aug 2003	0.86 (0.046)	1.53 (0.147)	−0.086 (0.092)	0.92	0.049	0.40
	237	25 Aug 2003	0.93 (0.048)	1.81 (0.146)	−0.21 (0.093)	0.94	0.043	1.13

Calibration and validation points pooled in 2001 ($n = 25$). All regression models are significant at $\alpha = 0.01$ level of significance and all 2002 and 2003 models are based on 12 calibration points except as noted. Numbers in parentheses represent standard errors (S.E.). The statistic I_{MS} is the Moran test for spatial independence (as defined in Eq. (4)). In 2001, off-the-shelf digital camera red wavebands were used in analysis. The NDVI calibrated to reflectance was used in 2002 and 2003. Height, width, and RMSE values for these attributes are in units of meters.

[a] One outlier deleted (see text).
[b] Calibration data $n = 11$.
[c] Significant at $\alpha = 0.05$ level.
[d] $p = 0.034$.
[e] $p = 0.026$.

The three *F*-tests used to assess model validity were defined as follows. First, note that the pooled **y** and **X** data can be partitioned as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \tag{5}$$

where the subscripts represent the calibration and validation data and regression matrices, respectively. Given this partition, the composite model *F*-test was performed by fitting:

$$\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon \tag{6}$$

and then testing if $\beta_1 = \beta_2$ (Cook and Weisberg, 1999). The joint prediction *F*-test was performed by first estimating Eq. (2) using just the calibration data, next calculating the joint set of prediction errors across the validation sites as

$$\mathbf{e}_2 = \mathbf{y}_2 - \hat{\beta}_1\mathbf{X}_2 \tag{7}$$

and then by computing the statistic:

$$F_1 = \mathbf{e}_2{}'\boldsymbol{\Omega}^{-1}\frac{\mathbf{e}_2}{s_1^2} \quad \text{where} \quad \boldsymbol{\Omega} = (\mathbf{I} + \mathbf{X}_2(\mathbf{X}_1{}'\mathbf{X}_1)^{-1}\mathbf{X}_2{}') \tag{8}$$

Note that F_1 follows a central $F(q, m)$ distribution under the null hypothesis (i.e., that the fitted calibration model is correct) where q and m represent the number of new (validation) prediction sites and the (calibration) model degrees

of freedom, respectively, and s_1^2 represents the estimated calibration model mean square error (Rao and Toutenburg, 1999). Finally, the mean prediction t-test was performed by first calculating the average prediction error as

$$\bar{e} = \mathbf{a}'\mathbf{e}_2 \quad \text{where } \mathbf{a}' = \begin{bmatrix} \frac{1}{q} & \dots & \frac{1}{q} \end{bmatrix} \tag{9}$$

and then computing the statistic:

$$t_1 = \frac{\bar{e}}{(s_1\sqrt{h})} \quad \text{where } h = \left[\left(\frac{1}{q}\right) + (\mathbf{a}'\mathbf{X}_2(\mathbf{X}_1{}'\mathbf{X}_1)^{-1}\mathbf{X}_2{}'\mathbf{a})\right] \tag{10}$$

Note that t_1 follows a central t distribution (with m degrees of freedom) under the null hypothesis, where s_1 is defined as before in Eq. (8) (Rao and Toutenburg, 1999).

Intuitively, the composite model F-test represents a test for non-equivalent intercepts and slopes across the partitioned calibration and prediction (validation) sites. In contrast, the joint prediction F-test assesses the ability of the regression model (fit using the calibration data only) to make unbiased predictions at all new validation sites, and simultaneously tests if these prediction errors are within the specified tolerance (precision) of the model. The mean prediction t-test follows from the joint prediction F-test, and hence assesses the ability of the regression model to make an unbiased prediction of the average value across the q validation sites.

We interpreted an I_{MS} test statistic greater than 1.65 to indicate statistically significant spatial correlation in the residual errors (based on a one-sided value for which a standardized normal z-score is 0.05). Likewise, we interpreted an F-test and/or t-test with a p-value <0.05 to indicate a violation of model validity (which we assume is due to sampling design induced bias, rather than model miss-specification). Note that the F- and t-test assume that the residual errors are normally distributed and jointly spatially independent. These assumptions were assessed using distribution tests and normal probability plots of the residuals (Cook and Weisberg, 1999), along with the results of the Moran test (for spatial correlation). The Moran score was calculated using the ESAP-calibrate software program (Lesch et al., 2000).

The ESAP-RSSD software attempts to maximize the probability of generating a sampling design that will yield spatially uncorrelated regression model residuals, but does not remove residual spatial correlation per se. It selects sites with lower probability of significant correlations by spreading out the sample locations within the confines of the field of interest. However, if the true residual spatial correlation structure is greater than the average between sample and site distance, then the observed residuals will tend to still exhibit spatial dependence. In general, the Moran test (or some similar test statistic) should always be used to validate the residual spatial independence assumption. Regression models, which exhibit a significant (residual) spatial correlation structure, are in theory sub-optimal; the model mean square error estimate and/or site-specific prediction intervals will typically be biased (i.e., under-estimated). Additionally, the corresponding parameter test statistics will tend to be anti-conservative (Cressie, 1991; Rao and Toutenburg, 1999), that is, the associated p-values will tend to be too small.

In 2002 and 2003, the validation step was not performed. Thus, only the first five steps used for the 2001 data were implemented: (1) remove outliers, (2) check residual error normality, (3) determine significance of correlation models for calibration sites, (4) produce regression statistics, and (5) calculate the Moran score. The root mean square error (RMSE) is an estimate of the standard deviation of the regression model (Neter et al., 1990). It was used here as a measure of the accuracy of the regression models in providing useful plant attribute values across the field.

3. Results

Crop width data for DOY 226 in 2001 equal to or greater than 100 cm were recorded as 100 cm causing truncated data, which violates the assumptions of regression modeling and, therefore, could not be modeled (Table 2). When the two outlier points for height (DOY 198 and 226) were removed, these regressions became significant at the $\alpha = 0.01$ level. All of the Moran (I_{MS}) test statistics associated with the 2001 regression models were non-significant and the residual errors appeared to be normally distributed. Thus, all regression models in 2001 satisfied the modeling assumptions required to perform the composite model and prediction tests necessary for validation of the sampling design.

If the regression models developed using the directed sampling approach are unbiased, then these equations should be able to predict the crop height and width measurements at all non-sampled locations in an unbiased manner. Additionally, the precision of these predictions should be within the specified precision of the fitted models. Thus, the

Table 3
Composite model and joint prediction *F*-tests and mean prediction *t*-test results for 2001 imagery vs. crop data

DOY (2001)	Canopy variable	Composite model *F*-score	Joint prediction *F*-score	Mean prediction *t*-score
198	Height[a]	3.03 (0.071)	0.92 (0.566)	−2.20 (0.053)
207	Height	1.25 (0.308)	0.49 (0.884)	−1.11 (0.293)
226	Height[a]	0.43 (0.657)	1.66 (0.226)	−0.06 (0.951)
198	Width	3.41 (0.052)	1.70 (0.204)	*−2.49 (0.032)*
207	Width	0.88 (0.429)	0.57 (0.828)	−0.41 (0.692)
226	Width[b]	–	–	–

Levels of significance (*p*) values are in parentheses. Test results that were significant below the $\alpha = 0.05$ level are shown in italics.
[a] One outlier deleted and removed (see text).
[b] Data truncated, violating assumptions of regression modeling.

composite and joint prediction *F*-tests and mean prediction *t*-tests performed in 2001 should all yield non-significant test scores (Table 3). Only the mean prediction *t*-score test for DOY 198 width appeared significant below the 0.05 level ($p = 0.032$). Furthermore, none of the 15 tests were significant at the 0.01 level. In Fig. 3, the validation site data are overlaid but are not included in the regression model. The inclusion of the validation points within the 95% confidence envelope for the calibration data offers further evidence that these two data sets described the same relationship. These test results indicate that the directed sampling approach was successful at choosing calibration sampling locations that could in turn be used for estimating unbiased prediction equations (i.e., regression models) for 2001. It should be clearly noted that these tests assess the ability of the ESAP response surface technique to produce a valid model and

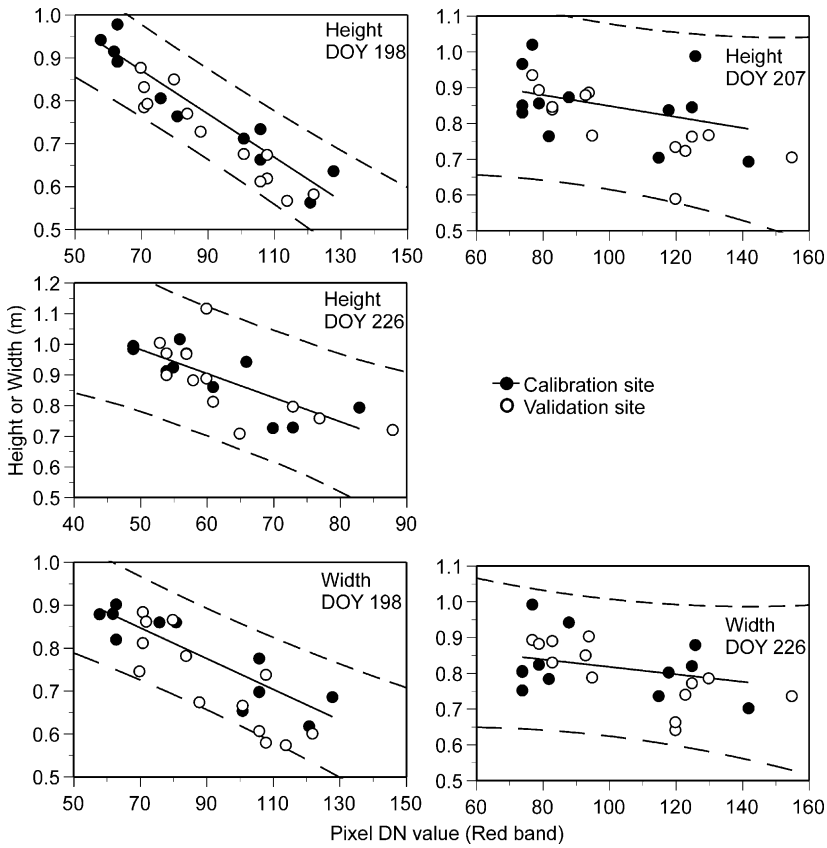

Fig. 3. Regressions of 2001 calibration vs. validation data (imagery pixel values) for crop height and width. Validation site data are overlaid but are not part of the regression. Crop data collected at the calibration and validation sites fall within the same data ranges, representing the same relationships. Dotted lines are 95% confidence intervals.

not to test the ability of the resulting regressions to predict the ground values and create useful maps of the variables of interest (height and width) as measured by r^2 or RMSE. This was done as a separate step for all 3 years.

The NDVI was not calculated in 2001 because mechanical defects prevented the use of a camera that would have provided near-infrared imagery. It is well established that spectral data in the red part of the spectrum and the NDVI are good estimates of plant features related to chlorophyll absorption, cover, height, etc. (Gausman, 1982; Jackson and Huete, 1991) during mid to late season (i.e., cover > 30%). Despite the different spectral measures of ground conditions (red band in 2001), the directed sampling approach will still be valid for the 2002 and 2003 NDVI data provided the response function describing the imagery/ground attribute relationships is the same (i.e., in this case, linear). The use of this response surface approach should be valid for any similar linear relationship. However, the ability of the model to predict ground attributes will depend on the variability of the data as represented by the RMSE. It is possible for this directed sampling (or any modeling) approach to be statistically valid but the predictive equations not very useful due to large amounts of scatter in the data.

Since the basic linear model was validated in 2001, validation data were not collected in 2002 and 2003. Co-located ground and imagery data in 2002 and 2003 and calibration locations in 2001 were used to develop regression models to convert (calibrate) imagery pixel values to estimates of the crop attributes (Table 2). All regression models but one were statistically significant at $\alpha = 0.05$ (Table 2) and most were significant at $\alpha = 0.01$. It is unknown why the DOY 192, 2002 date was not significant. Overall, the RMSE were acceptable, showing that this sampling method can predict height and width to within about 0.08 m with only 12 sampled locations.

Once regression statistics are calculated for the pixel values versus crop attribute (Table 2), predictive models can be generated and applied to the covariate data to produce maps of the crop attribute of interest (Fig. 4a and c). The ESAP-RSSD module outputs a 'survey' file that, along with a file containing the calibration data (width and height) at the sampled locations, can be input to the ESAP-calibrate module to produce the regression models, the Moran test statistic, and the prediction (and corresponding prediction variance) output files. Details of proper formatting and

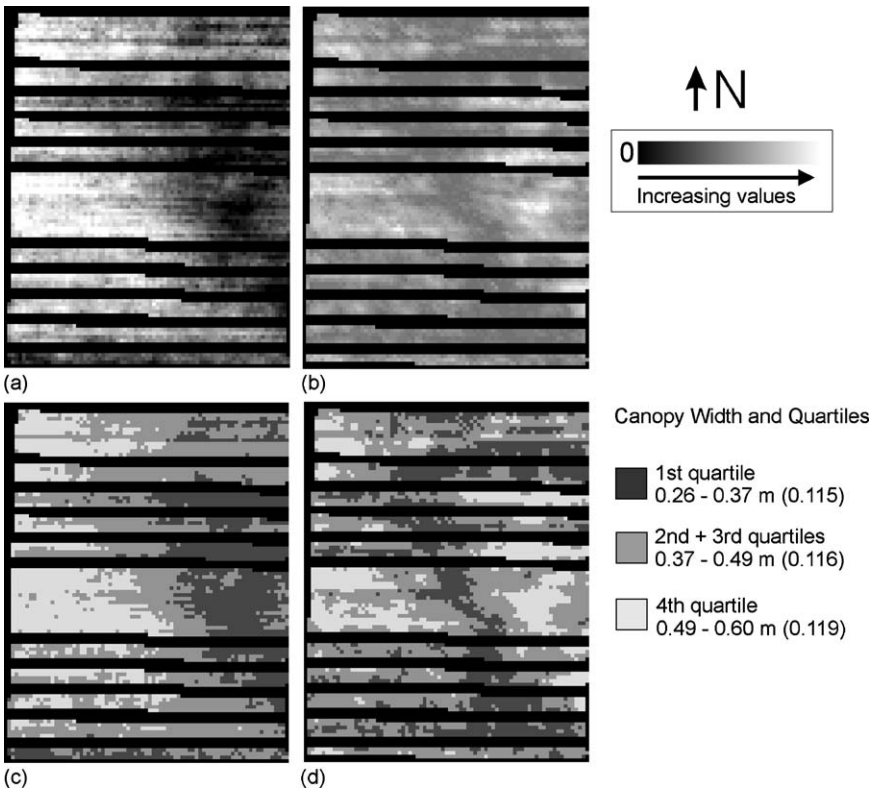

Fig. 4. Examples of predicted width (a) and standard deviation (b) maps for 2002, DOY 176 as output from the ESAP-calibrate module. The panels in (c) and (d) show the respective maps classified into quartiles with the 2nd and 3rd quartiles combined to simplify display. Each quartile range includes the mean standard deviation in parentheses.

processing steps can be found in Lesch et al. (2000). These model estimation and statistical prediction techniques can equivalently be carried out using most standard statistical software packages, if desired. An example of the predicted cotton crop width in 2002 and corresponding prediction standard deviation are shown in Fig. 4a and b. The corresponding classification maps (panels c and d) show the quartile ranges of the values and where they occur.

4. Discussion

Radiometrically calibrated or uncalibrated wide or narrow band imagery can be used for directed sampling as long as there is a linear (or low-order quadratic) relationship between the crop attributes of interest and the spectral bands. When these equations are used to predict crop attributes at other locations in the field, they should not be extrapolated outside the range of values selected by the ESAP-RSSD software early in the season since NDVI ~ 0.1 represented bare soil and a linear extrapolation to the *y*-axis would show positive values. The fitted line in Fig. 5 shows the regression for the pooled DOY 176–207, 2002 data extrapolated showing the presumed response to NDVI = 0.1. A scaled NDVI that sets the bare soil value to zero could also be used.

It should be noted that there were occasions when the 3-pixel sampling buffer around a given pixel had to be shifted by one pixel if a sample location fell near a skip row (and therefore the edge of a mask). It is possible that coordinate data were in error, since differential GPS data can contain errors of a few meters, although this was minimized due to meticulous geo-registration. The most likely reason for data scatter and therefore poorer correlations was that the small ground sampling areas may not have represented the 3 by 3 pixel areas well. It is important that the areas represented by ground sampling and pixel averaging are similar to reduce the chances of scale differences and data scatter (Atkinson and Curran, 1995).

It was noticed that the ESAP-RSSD software tended to be "attracted" to some points with more extreme values, such as bare soil (skips and missing plants) as well as field edges. Johnson et al. (2005) also reported this tendency of ESAP-RSSD to choose extreme EC_a values. By design, ESAP selects levels in the covariate data spanning the range of input values (Fig. 6), thus it will necessarily select some values near the extremes of the data set. The software also separates the locations apart as much as possible so points have a tendency to be located at the edges of fields or masked units within fields. Of the 120 calibration sites selected for the 10 dates, 48 were edge pixels. In the authors' experience, images should be carefully masked to remove certain features in the data such as obvious skips in the field and field edges. For example, the imagery could be filtered based on a given NDVI threshold (e.g., 0.2), removing soil-dominated pixels. If the imagery is of very high resolution, finer than the row structure of the crop, then variation is increased because the pixels contain information about individual plants, which may not be representative of larger areas in a field. Thus, imagery should also be re-sampled to a scale that is representative of the management question or ground sampling scale. For example, a 10-m pixel size might not be appropriate for a ground sample that represents a 1–2 m area (Atkinson and Curran, 1995). This issue of variation may cause ESAP to choose less representative points

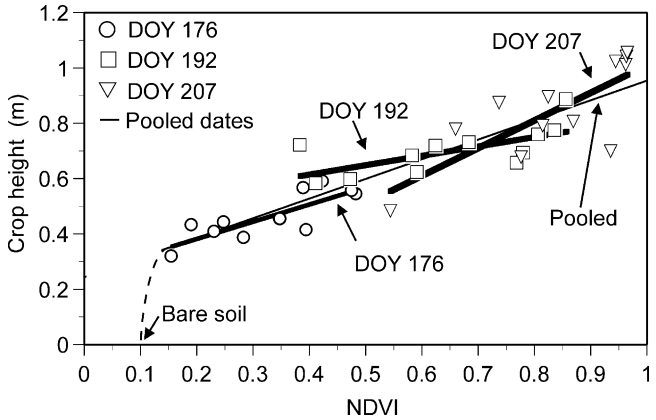

Fig. 5. NDVI from reflectance data vs. crop height for 2002 data plotted to show early and mid-season relationships. Other dates for height and width showed very similar relationships. The numbers represent day of year (DOY). The slope for pooled flights for all three dates is also shown. This shows that within each date slopes are linear but vary. The seasonal relationship is linear. The NDVI value of dry bare soil was about 0.1. As can be seen from the non-linear relationship early in the season, NDVI vs. crop response should not be extrapolated beyond the range of data collected.

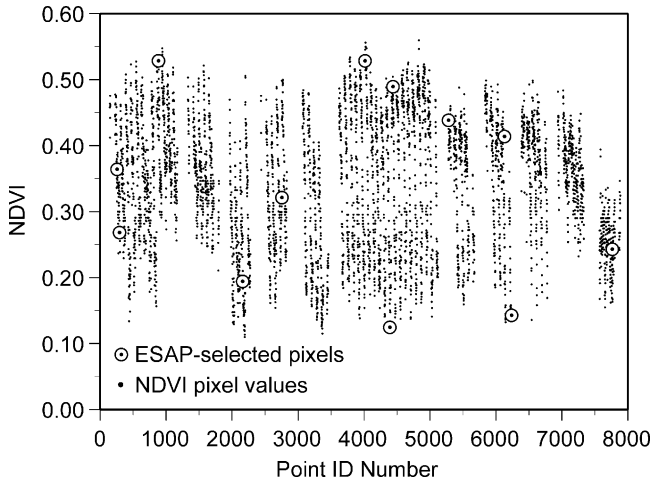

Fig. 6. Distribution of ESAP-selected points within the 2002, DOY 176 data set. The data points (pixels) were assigned sequential ID numbers and plotted on the *X*-axis vs. their respective NDVI values. Other dates showed similar distributions of ESAP-selected points across the pixel value data range.

for the question being addressed but is not unique to this method of choosing sampling locations. The scientific or management objective must always be matched to the scale of measurement. With these precautions in mind however, ESAP can choose representative data points for directed sampling.

In principle, predictive maps of the crop attribute of interest that are continuously variable across a field could be used directly by a producer or consultant or as input to a crop simulation model to extend the model spatially across a field. For example, canopy height and width are important parameters for estimation of evapotranspiration (Fitzgerald et al., 2003; Hunsaker et al., 2005). If the resolution of the continuously variable map or database is too fine for variable-rate management, then discrete management zones could be created based on thresholds or unsupervised classification (Fig. 4c). Another advantage of this directed sampling approach over classification or zone approaches (Chang et al., 2003) is that coordinates for site selection are provided. Thus, once a validated relationship is established between the remotely sensed imagery and crop attribute, application of this method by a crop consultant not trained in spectral classification would be relatively easy and finding sample locations less subjective than zone-based sampling techniques. Finally, compared to co-kriging and/or kriging with external drift, significantly fewer sampling locations are required for model calibration (thus reducing monitoring costs).

When using the directed sampling approach, the input covariate (i.e., remote) data can be an image or other spatial data set, including yield maps, digital elevation maps, and output from a model or GIS analysis, depending on the need of the analysis or management. This offers flexibility in terms of the data a farmer or consultant might have available. For example, plant height is one factor used in determining cotton growth regulator application rates and remotely sensed estimates of height have proven useful for variable-rate application of this input (Bethel et al., 2003; Lewis et al., 2003). Plant N status can also be estimated remotely and could be useful for variable nitrogen application maps (Bausch and Diker, 2001; Bronson et al., 2003; Fitzgerald et al., in press; Rodriguez et al., in press).

It should also be noted that the response surface technique can accommodate multiple inputs. The ESAP-RSSD software, as currently written, can accommodate two inputs from which sampling designs are generated. For example, two or more geo-referenced EC_a signals are commonly recorded during soil salinity surveys to help improved the accuracy of depth-specific salinity predictions (Rhoades et al., 1999). This could provide the opportunity to input two remotely sensed wavebands or indices for estimation of crop attributes, providing the relationship to the attributes of interest can be expressed as a (multiple) linear regression model.

Finally, systems-level research and management are hampered by a lack of tools to provided scale-bridging or scale-independent means of collecting and processing data. Although not tested here directly, it is hypothesized that if the relationship between the primary and covariate variables is of a low-order quadratic then this response surface sampling design procedure could be used at multiple scales of resolution from row to region. This would provide expanded opportunities for on farm studies and systems-level analysis using various types of geospatial inputs (Johnson et al., 2005).

5. Summary and conclusions

This study tested the ability of the ESAP-RSSD software to direct ground sampling using aerial imagery for purposes of predicting specific cotton crop attributes; i.e., height and width, from suitably calibrated regression equations. Although developed for soil salinity estimation, the response surface theory underlying the ESAP-RSSD software allows for the input of any geospatial remotely sensed data and maximizes the probability of generating a sampling design that will produce spatially uncorrelated regression model residuals. Relatively few locations are thus required (to estimate ordinary regression models) and maps of easily acquired crop attributes can be produced within about 36 h. The imagery for a field is processed within hours after a flight and ground sampling can occur the following morning. By afternoon, maps or digital databases can be output describing the spatial distribution of width, height or another easily recordable crop attribute.

The following outline summarizes the steps for using ESAP-RSSD with remotely sensed data to produce maps of crop attributes:

- Image processing:
  (1) Calibrate imagery to reflectance (optional);
  (2) Georegister imagery;
  (3) Mask undesirable areas based on thresholds or other methods;
  (4) Export as text file for input to ESAP-RSSD (requires three columns, *X* and *Y* coordinates in UTM and pixel value);

- ESAP-RSSD:
  (5) Remove outliers (if necessary) from imported pixel values using built-in routines and view histograms for data normality;
  (6) Generate sampling design and output to text files;

- GPS and ground data:
  (7) Load the final points to differential GPS unit;
  (8) Navigate to locations and collect field samples;

- ESAP-calibrate (or statistical software):
  (9) Process ground samples and organize data;
  (10) Perform regression analyses (prediction) and calculate Moran I_{MS} using ESAP-calibrate or another statistical package;

- Mapping only:
  (11) Input 'survey' file from ESAP-RSSD and ground data to ESAP-calibrate to develop prediction map;
  (12) Output prediction and variance data to external program (such as ENVI) to produce predictive and, optionally, standard deviation maps;

- Validation and Calibration:
  (13) If validating data (as done in 2001) then perform *F*-tests (composite, mean prediction, and joint prediction) using statistical software;
  (14) If Moran score and *F*-tests are not significant then create maps of the crop parameters of interest following steps 11 and 12.

Despite the use of different spectral wavebands and scatter in the data, all imagery-derived data were found to be significantly correlated to both crop height and width across multiple dates in three seasons of cotton. Measures that might be more useful to a producer or consultant, such as nitrogen status or evapotranspiration, could be used if the data satisfy the regression modeling assumptions. This approach has the advantage of supplying the farm manager with a small number of unambiguous sample locations that can be sampled quickly and processed for precision farming

applications. In practice, validation sites will not be needed for every date that data are collected if a reasonably strong *a priori* linear (or low-order quadratic) relationship can be shown to exist between the primary and covariate data from previous experimentation.

References

Atkinson, P.M., 2002. Spatial statistics. In: Stein, A., van der Meer, F., Gorte, B. (Eds.), Spatial Statistics for Remote Sensing. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 57–81.

Atkinson, P.M., Curran, P.J., 1995. Defining an optimal size of support for remote-sensing investigations. IEEE Trans. Geosci. Remote Sens. 33, 768–776.

Atkinson, P.M., Curran, P.J., 1997. Choosing an appropriate spatial resolution for remote sensing investigations. Photogramm. Eng. Remote Sens. 63, 1345–1351.

Bausch, W.C., Diker, K., 2001. Innovative remote sensing techniques to increase nitrogen use efficiency of corn. Commun. Soil Sci. Plant Anal. 32, 1371–1390.

Bethel, M., Gress, T., White, S., Johnson, J., Sheely, T., Roberts, B., Gat, N., Paggi, M., Groenenberg, N., 2003. Image-based, variable rate plant growth regulator application in cotton at Sheely Farms in California. In: Proceedings of the Beltwide Cotton Conference, The Cotton Foundation, Nashville, TN, USA, 6–10 January, pp. 1755–1771.

Bronson, K.F., Chua, T.T., Booker, J.D., Keeling, J.W., Lascano, R.J., 2003. In-season nitrogen status sensing in irrigated cotton. Part II. Leaf nitrogen and biomass. Soil Sci. Soc. Am. 67, 1439–1448.

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. 62, 241–252.

Chang, J.Y., Clay, D.E., Carlson, C.G., Clay, S.A., Malo, D.D., Berg, R., Kleinjan, J., Wiebold, W., 2003. Different techniques to identify management zones impact nitrogen and phosphorus sampling variability. Agron. J. 95, 1550–1559.

Cook, R.D., Weisberg, S., 1999. Applied Regression Including Computing and Graphics. John Wiley, New York, NY, USA, pp. 122–125.

Corwin, D.L., Plant, R.E. (Eds.), 2005. Applications of Apparent Soil Electrical Conductivity in Precision Agriculture, vol. 46. Computers and Electronics in Agriculture, pp. 1–3.

Cressie, N.A.C., 1991. Statistics for Spatial Data. John Wiley, New York, NY, USA.

Fitzgerald, G.J., Hunsaker, D.J., Barnes, E.M., Clark, T.R., Lesch, S.M., Roth, R., Pinter Jr., P.J., 2003. Estimating cotton crop water use from multispectral aerial imagery. In: Proceedings of the International Irrigation Association, The Irrigation Association, San Diego, CA, USA, November, pp. 138–148.

Fitzgerald, G.J., Rodriguez, D., Christensen, L.K., Belford, R., Sadras, V., Clarke, T., in press. Spectral and thermal sensing for nitrogen and water status in two wheat agroecosystems. Prec. Agric.

Gausman, H.W., 1982. Visible light reflectance, transmittance, and absorptance of differently pigmented cotton leaves. Remote Sens. Environ. 13, 233–238.

Haining, R., 1990. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press, Cambridge, England, pp. 146–147.

Hunsaker, D.J., Barnes, E.M., Clarke, T.R., Fitzgerald, G.J., Pinter, P.J., 2005. Cotton irrigation scheduling using remotely sensed and FAO-56 basal crop coefficients. Trans. ASAE 48, 1395–1407.

Huete, A., 1988. Soil-adjusted vegetation index (Savi). Remote Sens. Environ. 25, 295–309.

Jackson, R.D., Huete, A.R., 1991. Interpreting vegetation indices. Prev. Vet. Med. 11, 185–200.

Johnson, C.K., Eskridge, K.M., Corwin, D.L., 2005. Apparent soil electrical conductivity: applications for designing and evaluating field-scale experiments. Comput. Electron. Agric. 46, 181–202.

Lesch, S.M., Strauss, D.J., Rhoades, J.D., 1995a. Spatial prediction of soil salinity using electromagnetic induction techniques. Part 1. Statistical prediction models: a comparison of multiple linear regression and co-kriging. Water Resources Res. 31, 373–386.

Lesch, S.M., Strauss, D.J., Rhoades, J.D., 1995b. Spatial prediction of soil salinity using electromagnetic induction techniques. Part 2. An efficient spatial sampling algorithm suitable for multiple linear regression model identification and estimation. Water Resources Res. 31, 387–398.

Lesch, S.M., Rhoades, J.D., Corwin, D.L., 2000. ESAP-95 Version 2.01R: User Manual and Tutorial Guide. Research Rpt. 146, USDA-ARS, George E. Brown, Jr., Salinity Laboratory, Riverside, CA, USA.

Lesch, S.M., 2005. Sensor-directed response surface sampling designs for characterizing spatial variation in soil properties. Comput. Electron. Agric. 46, 153–179.

Lesch, S.M., Corwin, D.L., Robinson, D.A., 2005. Apparent soil electrical conductivity mapping as an agricultural management tool in arid zone soils. Comput. Electron. Agric. 46, 351–378.

Lewis, D.J., Fridgen, J., Johnson, J., Hood, K., 2003. Image-based site-specific plant growth regulator (SSPGR) applications at Perthshire Farms. In: Proceedings of the Beltwide Cotton Conference, The Cotton Foundation, Nashville, TN, USA, January 6–10, pp. 1772–1786.

Moran, M.S., Clarke, T.R., Qi, J., Pinter Jr., P.J., 1996. MADMAC: a test of multispectral airborne imagery as a farm management tool. In: Proceedings of the 26th Symposium on Remote Sensing of the Environment, Canadian Remote Sensing Society, Vancouver, Canada, BC, 25–29 March, pp. 612–617.

Moran, M.S., Bryant, R., Clarke, T.R., Qi, J., 2001. Deployment and calibration of reference reflectance tarps for use with airborne cameras. Photogramm. Eng. Remote Sens. 67, 273–286.

Neter, J., Wasserman, W., Kutner, M.H., 1990. Applied Linear Statistical Models: Regression, Analysis of Variance and Experimental Design, 3rd ed. Irwin Publishers, Homewood, IL, USA, p. 50.

Post, D.F., Mack, C., Camp, P.D., Sulliman, A.S., 1988. Mapping and characterization of the soils on the University of Arizona, Maricopa Agricultural Center. In: Proceedings of Hydrology and Water Resources in Arizona and the Southwest, University of Arizona, Tucson, AZ, USA, pp. 49–60.

Rao, C.R., Toutenburg, H., 1999. Linear Models: Least Squares and Alternatives, 2nd ed. Springer, New York, NY, USA, pp. 197–201.

Rhoades, J.D., Chanduvi, F., Lesch, S.M., 1999. Soil salinity assessment: methods and interpretation of electrical conductivity measurements. FAO Irrigation & Drainage Paper 57, FAO, Rome, Italy, pp. 63–93.

Rodriguez, D., Fitzgerald, G.J., Belford, R., in press. Detection of nitrogen deficiency in wheat from spectral reflectance indices and basic crop eco-physiological concepts. Australian J. Agric. Res.

Schabenberger, O., Gotway, C., 2005. Statistical Methods for Spatial Data Analysis. CRC Press, New York, NY, USA, pp. 215–295.

Upton, G., Fingleton, B., 1985. Spatial Data Analysis by Example: Point Pattern and Quantitative Data, vol. 1. John Wiley, New York, NY, USA, pp. 337–338.